Jenner, Matthew and Afonso, José P. and Kohlhaas, Christoph and Karbaum, Petra and Frank, Sarah and Piel, Jörn and Oldham, Neil J. (2016) Acyl hydrolases from trans-AT polyketide synthases target acetyl units on acyl carrier proteins. Chemical Communications, 52 (30). pp. 5262-5265. ISSN 1364-548X

Journal Name

COMMUNICATION

Acyl Hydrolases from *trans*-AT Polyketide Synthases Target Acetyl Units on Acyl Carrier Proteins

Received 00th January 20xx,
Accepted 00th January 20xx

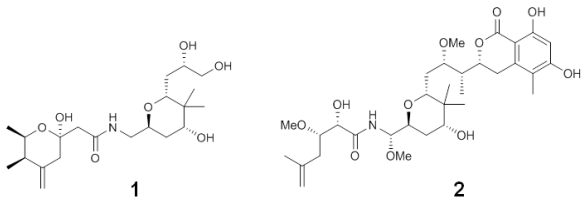
DOI: 10.1039/x0xx00000x

www.rsc.org/

Matthew Jenner[a,b], Jose P. Afonso[a], Christoph Kohlhaas[c], Petra Karbaum[c], Sarah Frank[c], Jörn Piel[c] and Neil J. Oldham[a]

Acyl hydrolase (AH) domains are a common feature of *trans*-AT PKSs. They have been hypothesised to perform a proofreading function by removing acyl chains from stalled sites. This study determines the substrate tolerance of the AH PedC for a range of acyl-ACPs. Clear preference towards short, linear acyl-ACPs is shown, with acetyl-ACP the best substrate. These results imply a more targeted housekeeping role for PedC: namely the removal of unwanted acetyl groups from ACP domains caused by erroneous transfer of acetyl-CoA, or possibly by decarboxylation of malonyl-ACP.

Modular type I polyketide synthases (PKSs) are responsible for the biosynthesis of a large array of complex, biologically active, natural products[1,2]. These molecular assembly lines utilise a strategy of covalent substrate attachment for high processivity. A phosphopantetheine chain, derived from co-enzyme A (CoA-SH), is employed to tether the substrate to acyl carrier protein (ACP) domains using a thioester bond. The so-called *trans*-AT PKSs constitute approximately 40% of all currently known bacterial multimodular PKSs[3], and represent a major, but poorly characterized, enzyme class of high relevance for drug discovery[4]. The most distinctive feature of modular *trans*-AT PKSs is the absence of acyltransferase (AT) domains integrated into each extension module. In these systems, the AT activity is supplied by *trans*-acting ATs often encoded within the gene cluster[5-7].



1 2

[a] *School of Chemistry, University of Nottingham, University Park, Nottingham, NG7 2RD (UK). E-mail: neil.oldham@nottingham.ac.uk*
[b] *Current address: Department of Chemistry, University of Warwick, Gibbet Hill, Coventry, CV4 7AL.*
[c] *Institute of Microbiology, Eidgenössische Technische Hochschule (ETH) Zurich, Wolfgang-Pauli-Strasse 10, 8083, Zurich (Switzerland); E-mail: jpiel.ethz.ch*

† Footnotes relating to the title and/or authors should appear here.
Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

However, a recently identified class of *trans*-acting AT-like domains have been shown to harbour hydrolytic activity towards acyl-thioesters[8]. These enzymes, subsequently classed as acyl hydrolases (AH), are distinctly different from AT domains at the sequence level and form a separate clade in a phylogenetic analysis of AT and AH domains from *trans*-AT PKSs (**Fig 1** and **S8**). AH domains have been proposed to act as proofreading enzymes for the PKS by hydrolysing stalled intermediates from the ACP domains of the biosynthetic assembly line[8,9]. In order to fulfil such a role, the AH domains would require a broad substrate specificity for a range of acyl chains. Herein we present a comprehensive substrate selectivity profile of the previously identified AH domain, PedC, from the pederin (**1**) PKS using a range of acyl-ACPs to mimic intermediates produced during pederin biosynthesis. Additionally, we investigate the role of a conserved Arg residue present in the active site of typical malonyl-incorperating AT domains, but absent from AH domains.

The PKS cluster encoding biosynthesis of the beetle toxin pederin (**1**), but belonging to an uncultivated bacterial symbiont[10], encodes two AT homologs; PedC and PedD[11]. Whilst the latter falls into the AT1 clade with other known malonyl transferases, PedC is located in the largely unstudied AT2 clade (**Fig 1** and **S8**)[8]. An *in vitro* functional analysis has shown that PedC catalyses the hydrolysis of acyl thioesters derived from both *N*-acetylcysteamine thioesters (SNACs) and acyl-ACPs, leading to its designation as a proof-reading enzyme that removes stalled polyketide intermediates[8]. A selection of short chain SNACs were previously employed to examine the specificity of PedC, with the 3-hydroxybutyryl, α,β-unsaturated and 4-methylpentanoic species among the better substrates.

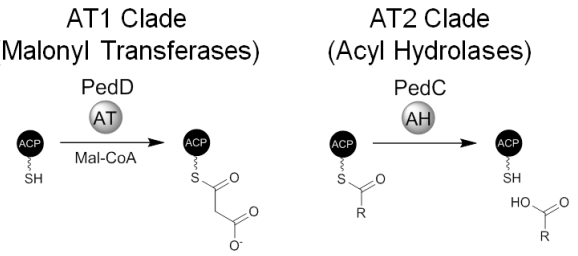


Fig 1.Phylogenetic analysis of all AT-like domains from *trans*-AT PKSs identifies two separate clades: AT1 comprised of standard malonyl transferase domains, and AT2 comprised of acyl hydrolase domains. The catalytic role of each domain type is shown.

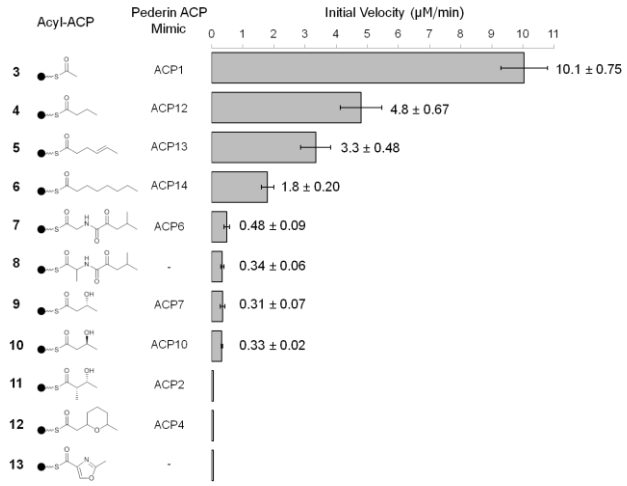


Fig 2 PedC-catalysed hydrolysis rates observed for acyl-ACPs **1-11**. Initial rate was calculated from pseudo-linear region of PedC hydrolysis plots. Errors in the mean of 3 repeats are given at a ± 2σ (95%) confidence interval. For acyl-ACPs **11-13** no hydrolysis greater than ~0.008 μM min^{-1} was observed over the time scale of the experiment (60 mins) compared to a control measurement.

PedC showed hydrolytic activity against both acetyl- and hexanoyl-ACP, but crucially, not malonyl-ACP which represents an essential precursor for polyketide biosynthesis[8].

Utilising available PsyA-ACP3 from the closely related psymberin (**2**) PKS from a bacterial sponge symbiont[11], a range of acyl-ACPs (**3 - 13**) was produced. The acyl moieties were designed to mimic the full range of functionality seen in the intermediates of pederin biosynthesis (**Fig S1**), thereby allowing us to test the ability of PedC to act as an efficient proofreading enzyme for the pederin PKS. The first aspect of PedC specificity probed was the acyl chain length. Here we employed straight chain acetyl, butyryl, hex-4-enoyl and octanoyl acyl-ACPs (**3 - 6**). Upon incubation of the acyl-ACPs with PedC and subsequent intact MS analysis, the appearance of the *holo*-ACP signal indicated hydrolysis of the acyl-ACP (**Fig S5**). Under the conditions employed, PedC successfully hydrolysed acetyl-ACP at an initial rate of 10.1 ± 0.75 μM min^{-1}, as measured from the pseudolinear region of hydrolysis time-course plot. Increasing the chain length to butyryl caused the rate to halve, 4.8 ± 0.67 μM min^{-1}. The hexenoyl- and octanoyl-ACPs were observed to have further decreased hydrolysis rates (**Fig 2**). The gradual decrease in hydrolytic rate with increased substrate chain length suggests that PedC is actually optimised for shorter chain intermediates.

The glycine-derived acyl-ACP (**7**) was used to approximate the pederin biosynthetic common β-amido intermediate at ACP6 (**Fig S1**). Limited hydrolysis was observed with (**7**) over the course of the experiment, with an initial rate of 0.48 ± 0.09 μM min^{-1}, which is approximately twenty-fold slower than the acetyl-ACP. An alanine derivative (**8**) was also tested with PedC to examine the effect of an α-methyl branch on the rate of hydrolysis. A small decrease in activity was observed for ACP (**8**), when compared to its non-α-methylated analogue (**7**). The (*R*) and (*S*) configured β-hydroxyl acyl-ACPs (**9** and **10**) were used to test whether PedC harbours any stereo-selectivity towards these acyl-chains. Both β-hydroxyl configurations are proposed to exist during pederin biosynthesis: occurring at ACP7 and ACP10 (**Fig S1**). The latter is believed to be a site of PedG oxygenase cleavage, resulting in the release of a pederin precursor[10]. Upon incubation with PedC, both (*R*) and (*S*) stereoisomers produced similar rates of hydrolysis: 0.31 ± 0.07 and 0.33 ± 0.02 μM min^{-1} respectively. These data show that no stereoselectivity is exhibited by PedC for these simple β-hydroxyl chains.

The α-methyl-β-hydroxyl acyl-ACP (**11**) is in the correct stereoconfiguration to act as a substrate mimic of the intermediate at ACP2 in pederin biosynthesis. Incubation of (**11**) with PedC yielded no measurable hydrolysis, implying that an α,β-branched substrate provides sufficient steric bulk to prevent hydrolysis under the conditions employed. The 6-methyltetrahydro-2*H*-pyran-ACP (**12**) was synthesised to represent the ACP4 intermediate in pederin biosynthesis. PedC was unable to hydrolyse this ACP species after 1 hr incubation. This result demonstrates that, in this assay, the pyran ring is too sterically demanding for the active site of PedC.

It has been proposed that the proofreading ability of PedC might be harnessed to elucidate the mechanistic steps in the biosynthesis of other biosynthetic pathways[9]. This requires PedC to be sufficiently promiscuous with regard to its specificity to hydrolyse unnatural intermediates from other clusters. Therefore, in addition to the acyl-ACP mimics of the pederin intermediates, we also tested oxazole-ACP (**13**), present as an intermediate of both rhizoxin and chivosazol biosynthesis[12,13]. No hydrolytic activity was observed towards the oxazole-ACP (**13**), suggesting that the five-membered heterocycle is not tolerated by the PedC active site. Previous work on the AH activity of PedC focussed principally on acyl-SNACs as substrate mimics[8] but limited work on acetyl- and hexanoyl-ACPs showed these to be successful substrate mimics at concentrations an order of magnitude lower than their SNAC equivalents. These observations, combined with the data presented here, indicate that acyl-ACPs represent more realistic and efficient substrates than the equivalent acyl-SNACs. This is in good agreement with previous studies on LovD, where hydrolytic release of the substrate was significantly improved in the ACP-bound form compared to the acyl-SNAC[14].

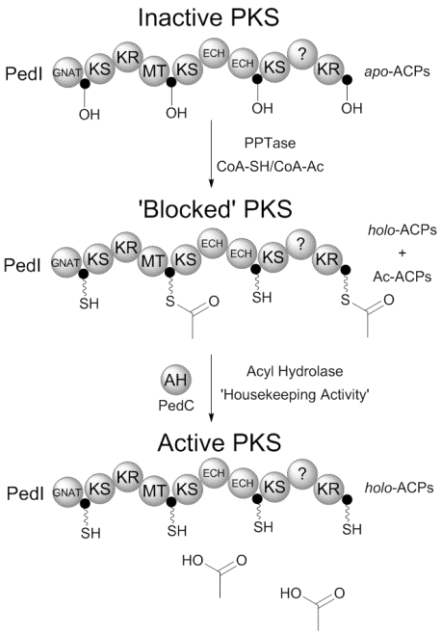


Fig 3 Proposed housekeeping role of PedC-like AH domains. A combination of high intracellular concentrations of Ac-CoA and high substrate tolerance of PPTases, results in loading of both Ac-CoA and CoA-SH onto the PKS assembly line. Acyl hydrolase 'housekeeping activity' is therefore required to hydrolyse acetyl units, ensuring an active biosynthetic cluster. Domain abbreviations: GNAT, GCN5-related N-acetyltransferase; KS, ketosynthase; KR, ketoreductase; MT, methyltransferase; ECH, enoyl-CoA hydratase; ?, domain of unknown function; •, ACP.

The acyl-ACP hydrolysis assay reveals that PedC has a strong preference for short, unbranched substrates: in particular the acetyl group. This result suggests that the hypothesised proofreading activity of PedC may be most effective in cases of inadvertent acetyl loading or unproductive decarboxylation of ACP-bound malonyl units, as previously suggested for type II TEs occurring in *cis*-AT PKS pathways[15-17]. The enzymes responsible for loading *apo*-ACPs with the phosphopantetheine chain from CoA-SH are the 4'-phosphopantetheine transferases (PPTases), which can be highly promiscuous with regard to ACP[18,19]. It has also been shown that several members of this class also transfer AcCoA (along with other derivatives) with high efficiency[20]. Given that – in anaerobic bacteria – the intracellular concentration of AcCoA can be up to 5-fold higher than that of CoA-SH[21], it is conceivable that ACPs are incorrectly loaded with AcCoA, as well as the desired CoA-SH.Loading of AcCoA would lead to effective 'blocking' the PKS assembly line at those points. By efficiently catalysing the hydrolysis of the acetyl-thioester bond, a 'housekeeping' role for AH domains would therefore be to ensure free thiols are available on the phosphopantetheine chain for substrate flux through the PKS (**Fig 3**).

A second potential mechanism for generating unwanted Ac-ACP is through decarboxylation of malonyl-ACP. This phenomenon has been previously proposed in the context of type II TEs occurring in *cis*-AT PKS pathways[17] and, although we are not aware of any direct evidence for general background decarboxylation, it could be a significant source of Ac-ACP. The half-life of bacterial proteins is typically quite high[22,23], therefore KS-catalysed, or indeed spontaneous decarboxylation of malonyl units may be an ongoing source of Ac-ACPs on the PKS. It might be hypothesised that PPTase-based transfer of Ac-CoA to the ACPs may only be an issue when the PKS is initially produced as *apo*-proteins. However the presence of ACP hydrolases in the cell, which catalyse the removal of the pantetheine moiety from ACPs[24-26], may compete with the loading process of PPTases allowing further acetyl units to be loaded.

The data presented thus far indicate that PedC is a hydrolytic enzyme with the ability to cleave short acyl intermediates from ACPs, whilst PedD is a malonyl-specific AT domain (**Fig 2 and Fig S7**). Despite such different catalytic roles, the domains are rather similar at the sequence level (~40% sequence similarity). In an attempt to rationalise the difference in catalytic activities of these domains, homology models of PedC and PedD were constructed using the CPHmodel server[27]. The models were subsequently aligned with the crystal structure of FabD, the AT domain from the *E. coli* fatty acid synthase (FAS), which was crystallised with a malonate unit attached to the active site serine[28]. This step allowed the same malonate unit to be modelled into the active site of the homology models. Analysis of the malonyl-complexed models for PedC and PedD revealed a strikingly obvious difference in the active site. The PedD and FabD structures both harbour an Arg residue at the back of the binding channel, which has been predicted to form a stabilising bidentate salt-bridge with the carboxyl region of the malonyl unit[28-31]. Furthermore, mutation of this Arg residue to a either Ala or Lys was found to abolish malonyl-specific transfer to the ACP in the fatty acid system[32]. In contrast to PedD, the same position in PedC is occupied by a neutral Gln residue, which would be unable to participate in an ionic interaction with the malonyl chain (**Fig 4**). In order to investigate whether this is a common feature of AH domains, a sequence alignment of all the AT and AH domains utilised in the polygenetic analysis performed (**Fig S8**). All domains predicted to be ATs (AT1 clade) harboured an Arg residue at this position. In contrast, all but one of the domains in the AT2 clade possessed a Gln residue in the equivalent position (**Fig 4**).

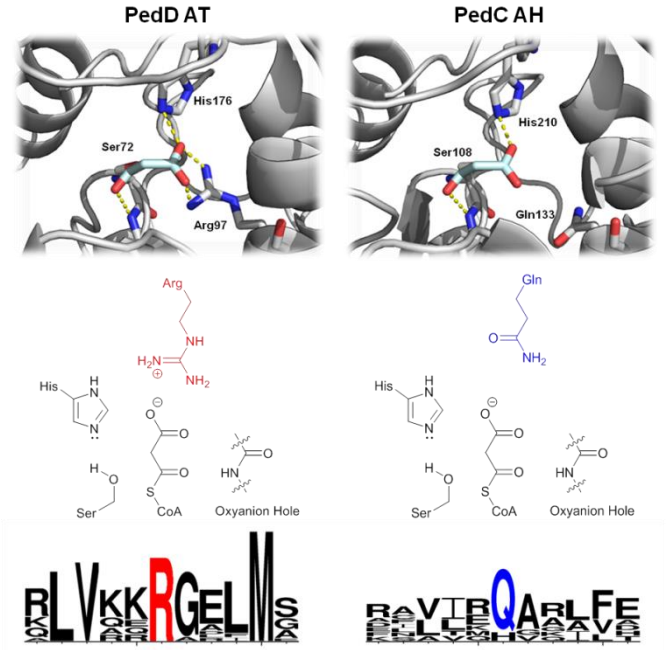


Fig 4 Homology models of PedD AT and PedC AH showing the arrangement of the active site residues around the malonate unit. The active site schematics simplify the interactions identified between the protein and the malonate. The major difference between the two domains is the presence of an Arg residue in PedD AT and a Gln residue in PedC AH. Also shown is a sequence logo representation of AT and AH alignments encompassing the Arg/Gln region.

Using this information, a PedD(R97Q) mutant was constructed to test the importance of the Arg residue towards malonyl-binding, and investigate the possibility that removal of this residue may switch the activity of PedD from an AT to a AH. PedD(R97Q) was incubated with increasing molar equivalents of malonyl-CoA to ascertain the ability of the mutant to be malonylated. However, even with 8 equivalents of malonyl-CoA, no malonyl-PedD(R97Q) was observed (**Fig S9**). This suggested that removal of the stabilising Arg residue (**Fig 4**) has effectively rendered the AT domain catalytically inactive as a malonyl-transferase. The lack of activity towards malonyl-CoA was similar to that found for PedC. In order to investigate any potential AH activity of PedD(R97Q), the mutant was incubated with acetyl-ACP at a 1:1 molar ratio. After 5 min, approx. 45% of the acetyl group was removed from the PPant thiol of the ACP, with WT PedD exhibiting no activity (**Fig S10**). This result indicated that the point mutation may have yielded hydrolase activity; however, leaving the reaction for longer (20 – 30 mins) resulted in no further deacylation. This suggested that following acyl transfer to PedD(R97Q), no hydrolysis of the acyl-enzyme intermediate occurred. Examination of the PedD(R97Q) MS signals indicated that the latter is correct, with approximately 45% of PedD(R97Q) observed to be acetylated, congruent with the levels of de-acetylation of the ACP (**Fig S10**). The comparable levels of acetylation observed on both PedD(R97Q) and ACP suggested an approximately balanced equilibrium of de-acetylation/re-acetylation had been established.

Although the Arg→Gln point mutation did not generate an enzyme with true acyl hydrolase activity, the ability of the domain to allow transfer of an acyl chain to the active site serine is different from the activity exhibited by WT PedD, which is unable to catalyse

either transfer or hydrolysis of an acyl chain. These results indicate that the point mutation has changed the activity of PedD towards that of PedC, with the structural factors dictating the ability to catalyse hydrolysis yet to be identified. It has been proposed that different oxyanion-hole loop orientations in the active site act as a differentiating factor between AT and AH activities[33].

In conclusion, we have successfully probed the substrate specificity of the PedC AH using a range of acyl-ACPs, revealing marked preference for short, linear acyl chains: in particular the acetyl group. This observation leads us to re-assess the principal role of AHs in PKS assembly lines as housekeeper enzymes, serving to hydrolyse erroneously loaded acetyl groups on ACPs. It should be pointed out that previous work on bacillaene and rhizoxin PKS possessing inactivated terminal thioesterase domains showed that short- and long-chain intermediates were hydrolysed off the PKS in cells[34-36]. This finding might be due to an additional proofreading activity present in these pathways, a different activity profile of the cognate AHs, or simply base-catalysed hydrolysis. Alternatively, PedC and its homologues could hydrolyse off the acyl chain with reasonable efficiency under conditions at which polyketide biosynthesis has completely stalled, as would be the case for the inactivated thioesterase variants. Comparative homology modelling and sequence analysis of AT and AH domains allowed identification of a critical Gln residue in active site of AH domains, where archetypal AT domains possess an Arg residue in the corresponding position, promoting malonyl-binding. Mutation of Arg→Gln in PedD AT confirmed the crucial role of this residue, abolishing malonyl-binding, but allowing acetyl transfer onto the active site of PedD(R97Q). However, other structural aspects are clearly involved in the hydrolysis of the acyl unit from the active site.

NJO thanks the Leverhulme Trust for a Research Grant (RPG-2012-578), and JP is grateful to the DFG for funding (PI 430/8-1).

Notes and references

1. G. M. Cragg and D. J. Newman, *Pure and Appl. Chem.,* 2005, **77**, 7-24.
2. C. T. Walsh, *Science.* 2004, **303**, 1805-1810.
3. R. V. O'Brien, R. W. Davis, C. Khosla and M. E. Hillenmeyer, *J. Antibiot.,* 2014, **67**, 89-97.
4. G. M. Cragg and D. J. Newman, *J. Nat. Prod.* 2007, **70**, 461-477.
5. J. Piel, *Nat Prod Rep.* 2010, **27**, 996-1047.
6. M. Till and P. R. Race, *Biotechnol. Lett.,* 2014, **36**, 877-888.
7. Y. Q. Cheng, G. L. Tang and B. Shen, *Proc. Natl. Acad. Sci. U. S. A.,* 2003, **100**, 3149-3154.
8. K. Jensen, H. Niederkrüeger, K. Zimmermann, A. L. Vagstad, J. Moldenhauer, N. Brendel, S. Frank, P. Pöplau, C. Kohlhaas, C. A. Townsend, M. Oldiges, C. Hertweck and J. Piel, *Chem. Biol.,* 2012, **19**, 329-339.
9. J. C. Kwan, E. W. Schmidt, *Chem. Biol.* 2012, **19**, 309-311.
10. J. Piel, *Proc. Natl. Acad. Sci. U. S. A.,* 2002, **99**, 14002-14007.
11. K. M. Fisch, C. Gurgui, N. Heycke, S. A. van der Sar, S. A. Anderson, V. L. Webb, S. Taudien, M. Platzer, B. K. Rubio, S. J. Robinson, P. Crews and J. Piel, *Nat. Chem. Bio.,* 2009, **5**, 494-501.
12. B. Kusebauch, N. Brendel, H. Kirchner, H-M. Dahse and C. Hertweck, *Chembiochem.* 2011, **12**, 2284-2288.
13. O. Perlova, K. Gerth, O. Kaiser, A. Hans and R. Müller, *J. Biotechnol.,* 2006, **121**, 174-191.
14. Xie X, Meehan MJ, Xu W, Dorrestein PC, Tang Y., *J Am. Chem. Soc.,* 2009, **131**, 8388-8389.
14. E. Yeh, R. M Kohli, S. D. Bruner and C.T. Walsh, *Chembiochem.,* 2004, **5**, 1290-1293.
15. D. Schwarzer, H. D. Mootz, U. Linne and M. A. Marahiel, *Proc. Natl. Acad. Sci. U. S. A.,* 2002, **99**, 14083-14088.
16. M. L. Heathcole, J. Staunton and P. F. Leadlay, *Chem. Biol.* 2001, **8**, 207-220.
17. A. K. Joshi, L. Zhang, V. S. Rangan, and S. Smith, *J. Biol. Chem.* 2003, **278**, 33142-33149.
18. C. T. Walsh, A. M. Gehring, P. H. Weinreb, L. E. N. Quadri and R. S. Flugel, *Curr. Opin. Chem. Biol.,* 1997, **1**, 309-315.
19. J. Beld, E. C. Sonnenschein, C. R. Vickery, J. P. Noel and M. D. Burkart, *Nat. Prod. Rep.,* 2014, **31**, 61-108.
20. S. Chohnan, H. Furukawa, T. Fujio, H. Nishihara and Y. Takamura, *Appl. Environ. Microbiol.,* 1997, **3**, 553-560.
21. A. L. Koch and H. R. Levy, *J. Biol. Chem.,* 1955, **217**, 947-957.
22. J. Mandelstam, *Biochem. J.,* 1958, **69**, 110-119.
23. P. R. Vagelos and A. R. Larrabes, *J. Biol. Chem.,* 1967, **242**, 1776-1781.
24. E. Murugan, R. Kong, H. Sun, F. Rao and Z-X. Liang, *Protein Expr. Purif.,* 2010, **71**, 132-138.
25. N. M. Kosa, K. M. Pham and M. D. Burkart, *Chem. Sci.,* 2014, **5**, 1179-1186.
26. M. Nielsen, C. Lundegaard, O. Lund and T. N. Petersen, *Nucleic Acids Res.,* 2010, **38**, W576-W581.
27. C. Oefner, H. Schulz, A. D'Arcy, and G. E. Dale, *Acta Crystallogr. Sect. D-Biol. Crystallogr.,* 2006, **62**, 613-618.
28. Y. Y. Tang, C. Y. Kim, I. I. Mathews, D. E. Cane and C. Khosla, *Proc. Natl. Acad. Sci. U. S. A.,* 2006, **103**, 11124-11129.
29. F. T. Wong, X. Jin, I. I. Mathews, D. E. Cane, and C. Khosla, *Biochemistry.,* 2011, **50**, 6539-6548.
30. A. T. Keatinge-Clay, A. A. Shelat, D. F. Savage, S. C. Tsai, L. J. Miercke, J. D. O'Connell, C. Khosla, R. M. Stroud, *Structure,* 2003, **11**, 147-154.
31. V. S. Rangan and S. Smith, *J. Biol. Chem.,* 1997, **272**, 11975-11978.
32. Y. Jiang, K. L. Morley, J. D. Schrag and R. J. Kazlauskas, *Chembiochem.* 2011, **12**, 768-776.
33. J. Moldenhauer, X. H. Chen, R. Borriss, J. Piel, *Angew. Chem. Int. Ed.,* 2007, **46**, 8195-8197.
34. J. Moldenhauer, D. C. G. Goetz, C. R. Albert, S. K. Bischof, K. Schneider, R. D. Süssmuth, M. Engeser, H. Gross, G. Bringmann and J. Piel, *Angew. Chem. Int. Ed.,* 2010, **49**, 1465-1467.
35. B. Kusebauch, B. Busch, K. Scherlach, M. Roth and C. Hertweck, *Angew. Chem. Int. Ed.,* 2009, **48**, 5001-5004.

